

Available online at www.sciencedirect.com**SciVerse ScienceDirect**

Procedia Environmental Sciences 11 (2011) 447 – 451

Procedia
Environmental Sciences

Research of Oracle Bone Inscriptions Ontology Construction Based on Relational Database

Xiong Jing, Liu Yong-ge , Gao Feng , Wang Ji-peng

School of Computer and Information Engineering

Oracle Bone Inscriptions Information Processing Laboratory

Anyang Normal University, Anyang, China

jingxiong125@gmail.com, ay_liuyongge@163.com ,gaof97@163.com, wangjipeng2002@yahoo.com.cn

Abstract

According to the difficulty of Oracle Bone Inscriptions(OBI) knowledge sharing, a scenario of OBI ontology construction is proposed. Because of the hard work of ontology building, a transformation method from existing OBI database to ontology is discussed. Firstly, analysed the similarity between entity-relationship model and RDF model. Secondly, introduced the transformation ideas and process. Finally, implemented OBI ontology by using the transformation method. It is a semi-automatic method because there requires manual intervention during the process of transformation. Results show that the method can really reduce the burden of ontology researchers.

© 2011 Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

Selection and/or peer-review under responsibility of the Intelligent Information Technology Application Research Association.

Keywords: Oracle Bone Inscriptions; ontology construction; relational database; RDF triples

1.Introduction

OBI has more than 3,500 years' history. It record a wide range of social activities of the Shang Dynasty royal divination, including Wang Shi, agricultural, astronomical phenomena, good and bad, sacrifice, conquest, imperative, exchanges and marriage. Because of its very rich historical content, it has important historical value. OBI's in-depth research has a profound impact on language, history, archeology, social anthropology, ancient history and other disciplines of science and technology^[1]. But it is very difficult to identify and collate OBI. Due to the rapid development of computer information technology, it prives a very convenient conditions to research the ancient Chinese deeply. Therefore, using information science and technology to carry out digitization of OBI, is a very important way to promote and develop OBI research.

OBI is a high threshold in the study. Currently, OBI can be identified and translated by only a few persons. On the one hand, there are very few of OBI Experts, OBI professional training need a decade or two or even longer^[1]. On the other hand, OBI experts' identification and translation rely on the long-term academic study and experience. These experience knowledges are only stored in the minds of experts, and

effective sharing of knowledge can not be achieved. Even there are some relevant OBI publication, the traditional publication is not easy to be understood and handled by machines. How to use existing technology to achieve an efficient and effective knowledge sharing of OBI and make it easy to understand and deal with for computers?

Ontology as a shared and clear conceptual model of formalization can provide a solution to this problem. But ontology construction is a complicated process and requires a lot of manpower and time. So, a semi-automatic construction of OBI ontology based on the relational database is considered.

Currently, there are many studies and researches about the mapping between relational database and ontology and some of them have gained high achievement. Reference [2] analyzed the correspondent relationship between relational database and OWL ontology and designed a set of conversion rules in order to make use of the existing relational database to generate ontology automatically. Reference [3] proposed a novel automatic transform method from relational database to ontology. By the analysis of primary keys, attributes, foreign keys, integrity constraints of relational model and partial data. The method could construct ontology while conserving the information of the relational databases and fulfilling primary integration and classification. Reference [4] presented a formal approach and an automated tool for translating ER schemata into Web ontologies in the OWL. Following the predefined knowledge-preserving mapping rules from ER schema to OWL DL ontology, it could automatically translates the schema into the ontology in both the abstract syntax and the RDF/XML syntax for OWL DL.

2. Entity-relationship model

The basic elements of entity-relationship are entities, relationships and values. An entity is a thing which can be distinctly identified. Entities can be classified into different entity sets. A relationship is an association among entities. A relationship set is a set of relationships of the same type. In an entity set $\{[e_1, e_2, \dots, e_n] | e_1 \in E_1, e_2 \in E_2, \dots, e_n \in E_n\}$, each tuple of entities $[e_1, e_2, \dots, e_n]$ is a relationship and relationship set R_i is a mathematical relation among n entities. The information about an entity or a relationship is obtained by observation or measurement, and is expressed by a set of attribute-value pairs [5].

The entity-relationship diagram is used for describing entities and relationships. An entity-relationship diagram is a specialized graphic that illustrates the interrelationships between entities in a database. In an entity-relationship diagram, a relationship set may be defined on only one entity set or more than two entity sets. There also may be more than one relationship set defined on given entity sets. There are three mappings between entities, 1:1, 1:n and m:n.

Fig.1 illustrates an example of entity-relationship diagram about students and courses.

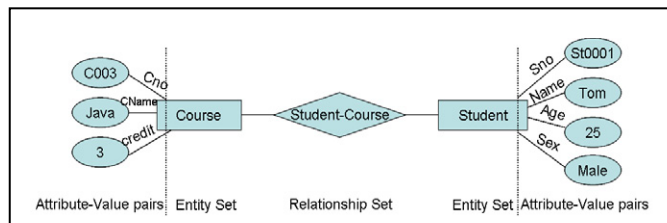


Figure 1. A simple entity-relationship diagram.

3.RDF model

The foundation of RDF is a model for representing named properties and property values. The RDF model draws on well-established principles from various data representation communities. RDF properties may be thought of as attributes of resources and in this sense correspond to traditional attribute-value pairs. RDF properties also represent relationships between resources and an RDF model can therefore resemble an entity-relationship diagram. Precisely speak, RDF is a member of the entity-relationship modelling family. So, RDF Schemas which are themselves instances of RDF data models are ER diagrams. In object-oriented design terminology, resources correspond to objects and properties correspond to instance variables.

The basic data model of RDF consists of three object types:

Resources All things being described by RDF expressions are called resources such as an element of a HTML, a entire Web page, a whole Web site at el.Even a printed book is a resource.

Properties A property is a specific aspect, characteristic, attribute, or relation used to describe a resource. Each property has a specific meaning, defines its permitted values, the types of resources it can describe, and its relationship with other properties.

Statements A specific resource together with a named property plus the value of that property for that resource is an RDF statement. The statements describe the resources in the form of subject-predicate-object expressions. These expressions are known as triples in RDF terminology. The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object. The object of a statement can be another resource or it can be a literal.

For example, one way to represent the notion “Tom chooses the course Java which the Cno is C003 and the credit is 3 ” in RDF is as the triple: a subject denoting "Tom", a predicate denoting "choose", and an object denoting "course". Fig.2 illustrates the statement using RDF graph.

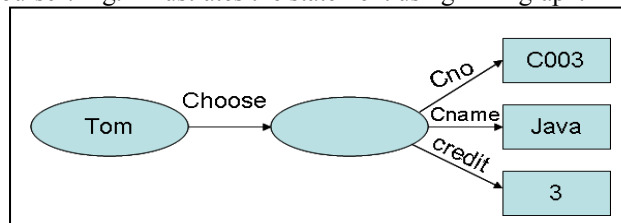


Figure 2. A simple RDF graph.

4.Relation-ontology transformation

From Fig.1 and Fig.2, it can be seen that there are similar elements and structure between entity-relationship model and RDF model. So we can realize the semi-automatic ontology building from relational database. Fig.3 illustrates the process of transform from relational database to ontology.

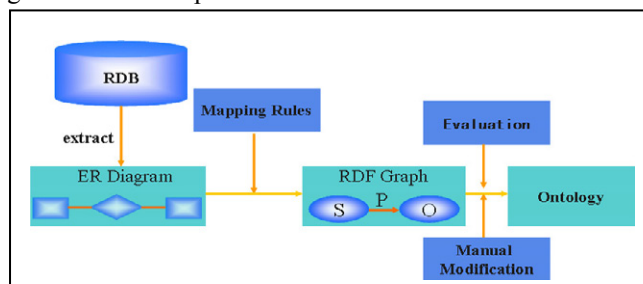


Figure 3. Transformation process of RDB to ontology.

The transformation process shown in Fig.3 are as follows:

Step 1: Extract the entity-relationship diagram from relational database by using reverse engineering.

Step 2: Transform the entity-relationship diagram to RDF graph by using some additional mapping rules(Here, we use ER Diagram and RDF Graph to illustrate the flow of transformation. What we really need are ER information and RDF triples). Each row in the database row describes a single entity, all of the same type. Each row will have a distinct URI because the row represents a distinct entity. Then use that URI as the namespace for all the identifiers in the database. Each row in the table says several things about its items. To represent this in RDF, each of these will be a property that will describe the entities. There will be one triple per cell in the table—that is, for n rows and c columns, there will be $n \times c$ triples. Each entity in the table can be represented in RDF by adding one tripe per row that specifies the type of the individual described by each row[6]. Mapping rules discussed in [7] are choosed as our rules during the transformation from RDB to RDF.

Step 3: Check and evaluate the RDF graph(triples) generated in Step 2. If there are some wrong expressions, correct them by using manual method.

Step 4: Create the ontology in RDF/XML format through RDF triples.

5.OBI Ontology automatic building based on relational database

Based on the transformation discussed in section IV. We built an OBI ontology automatically from OBI database. Fig.4 shows the table information in OBI database and Fig.5 shows some RDF triples corresponding to Fig.4.

number	OracleBoneWord	TraditionalChinese	RedefinitionWord	SimplifiedChinese
001162	炆	競		竞
001163	𩇛	競		竞
001164	𩇛	競		竞
001165	𩇛	競		竞
001166	𩇛	眾		众
001167	𩇛		𩇛	
001168	𩇛		𩇛	
001169	𩇛	鬥		斗
001170	𩇛	任		任
001171	𩇛	任		任

录: 1 共有记录数: 6199

Figure 4. Table information of OBI database.

Subject	Predicate	Object
obi:word1	obi: Word_Number	001162
obi:word1	obi:OBI_word	炆
obi:word1	obi:Tradition_Chinese	競
obi:word1	obi:Simplified_Chinese	竞
obi:word2	obi: Word_Number	001163
obi:word2	obi:OBI_word	𩇛
obi:word2	obi:Tradition_Chinese	競
obi:word2	obi:Simplified_Chinese	竞

Figure 5. Some RDF triples corresponding to Fig.4.

When we get the RDF triples, we can generate the ontology by some tools. Jena API[8] is used to create RDF files in our research. Some code fragments are as follows:

```

<obi:OBI rdf:about="&obi;OBI_word1"
  obi:OBI_word="炆"
  obi:Simplified_Chinese="竞"
  obi:Tradition_Chinese="競"
  obi:Word_Number="001162"/>
<obi:OBI rdf:about="&obi;OBI_word2"
  obi:OBI_word="藜"
  obi:Simplified_Chinese="竞"
  obi:Tradition_Chinese="競"
  obi:Word_Number="001163"/>
</rdf:RDF>

```

6. Conclusions Remarks

We have proposed an approach to semi-automatic transformation of relational databases to ontologies, where the quality of transformation is also considered. We used the method to implement OBI ontology construction. It can really lighten the researcher's burdens. In the future, we will extend the OBI database and consider OWL ontology building based on it. The correctness verification of the transformation from relational database to OWL ontologies is also an important work.

Acknowledgment

This research was supported by National Natural Science Foundation of China (No. 60875081).

References

- [1] Jiang Ming-hu, Natural Language Processing, BeiJing: Higher Education Press, 2006.
- [2] Chen He-ping, Hu Lu, and Chen Bin, et al, "Design and Implementation of Ontology Generator Based on Relational Database," Computer Engineering, 2009, vol 35(5), pp. 34-36.
- [3] Yu Xia, Liu Qiang, and Ye Dan, "Rule-based approach to transform relational database to ontology," Application Research of [1] Computers, 2008, vol 25(3), pp. 767-770.
- [4] Xu Zhuo-ming, Cao Xiao, Dong Yi-sheng, et al, "Formal approach and automated tool for translating ER schemata into OWL ontologies," Berlin: Springer Advances in Knowledge Discovery and Data mining: 8th Pacific-Asia conference (PAKDD 2004), May. 2004, pp. 464-475.
- [5] PPS Chen, "The entity-relationship model—toward a unified view of data," ACM Transactions on Database Systems (TODS), 1976, 1(1), pp. 9-36.
- [6] Allemang D, Hendler J A, Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL, Morgan Kaufmann, 2008.
- [7] Shen, G., Huang, Z., Zhu, X., et al, "Research on the Rules of Mapping from Relational Model to OWL," in: Proceedings of the Workshop on OWL: Experiences and Directions, 2006, Vol. 216.
- [8] B. McBride, "Jena: Implementing the RDF model and syntax specification," in: Semantic Web Workshop, WWW, 2001.